DOCUMENT RESUME

ED 073 051 SP 005 994

AUTHOR Derhammer, John L.; Cormier, William H.
TITLE The Effects of Training on Increased Use of and Correlation Between Higher Order Classroom and Test Questioning Behavior.
PUB DATE Apr 72
NOTE 23p.; Paper presented at the annual meeting of the Educational Research Association, Chicago, Illinois, April 5, 1972

EDRS PRICE MF-$0.65 HC-$3.29
DESCRIPTORS *Feedback; *Questioning Techniques; *Teacher Behavior; *Teacher Education; *Teaching Techniques

ABSTRACT

This controlled study investigaged the effects of training, with and without feedback, on increasing teachers' higher-order classroom and test questions and improving the correlation between these two kinds of questioning behavior. A Questions Analyzer was used to obtain the data from 12 elementary school teachers in Tennessee. Classroom and test questioning behaviors were analyzed using analysis of variance and coefficient of correlations. Results indicated that training significantly increased the incidence of, and correlation between, higher-order classroom and test questioning behaviors in both trained groups. The findings emphasize the importance of training teachers in the use of higher-order questions. (A 20-item bibliography is included.) (Author)

ED 073051

The Effects of Training on Increased Use of And Correlation Between Higher Order Classroom and Test Questioning Behavior

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY

John L. Derhammer
Georgia State University
Fort Benning Center
Columbus, Georgia

William H. Cormier
University of Tennessee
Knoxville, Tennessee

SP 005 994

A paper presented at the 1972 American Educational Research Association Annual Meeting - Chicago, Illinois, April 5, 1972

The Effects of Training on Increased Use of and Correlation Between Higher Order Classroom and Test Questioning Behavior

John L. Derhammer
Georgia State University, at Columbus

and

William H. Cormier
The University of Tennessee

The controlled study investigated the effects of training, with and without feedback, on increasing teachers' higher order classroom and test questions and improving the correlation between these two kinds of questioning behavior. A Questions Analyzer was used to obtain the data from twelve elementary school teachers in Tennessee. Classroom and test questioning behavior were analyzed using analysis of variance and coefficient of correlations. Results indicated that training significantly increased the incidence of, and correlation between, higher order classroom and test questioning behavior in both trained groups. The findings emphasize the importance of training teachers in the use of higher order questions.

Even the most severe critics of education would probably agree that teachers sincerely desire that their students, upon completion of a course of study, will be able to transfer "classroom learning" to other environmental settings. Despite the best intentions of teachers a long standing weakness in many educational programs has been that of bridging the gap between the classroom and the "real world."

It has been charged by some educators that teachers actually contribute to this weakness by stressing the mere memorization of facts and by failing to provide laboratory controlled situations where students can practice the mental activities required in critical thinking (19, 7, 6).

Research has indicated that approximately 90 percent of teachers' questions in the classroom require students only to memorize and recite facts while teacher made tests reveal an even higher percentage of factual recall questions (9, 5, 15).

Although it has been fairly well established that teachers ask few higher order questions---practically all reported findings are in agreement even as to the percentages---some confusion exists as to just what training methods will work to increase the teachers' use of higher order questions.

Several studies, all of which contain possible weaknesses of design, provide conflicting reports regarding the effects of feedback training or a combination of these two variables and their various levels of application. Mittlestadt (13), Cunningham (4), Schreiber (17), Meehan (12), Cornell (3), Konetski (10), and Claus (2) all reported that various training methods altered questioning behavior but that none appeared to be particularly effective in comparison with others. On the other hand Manson (11) reported that his training methods failed to increase higher order questioning practices. The confusion exists in this area

because of the inclusion of two independent variables (or more in some instances) ---training and/or feedback---and, as a consequence, reports that one or the other or both are effective or ineffective. Most of the above cited research indicates the effectiveness (or lack of effectiveness) of one or the other variable without providing adequate control. Hopefully the present research will shed more light on this aspect of the problem.

Bloom (1), Guilford (7) and Sanders (16) have provided practical suggestions for educators who desire to train students in critical thinking (problem solving) through the implementation of "higher order" (more than memorization of facts) questions. The assumption implied by the above researchers is that teachers can be trained to develop problem solving skills in their students through inservice or preservice exposure to a variety of question forms that will produce the desired mental activity for solving a particular problem. Since previous research has indicated that teachers use few higher order questions it would appear that this assumption is open to question. The proposed training of teachers might present some difficulties in itself.

With these thoughts in mind the present study attempted (1) to increase the number of higher order questions asked by inservice teachers both in the classroom and on tests and (2) to increase the correlation between higher order classroom and test questioning behavior.

The following terms, mainly from Sanders' (16) are defined to provide clarification of meaning.

Higher order questions: Questions asking the student to do more than repeat materials he has memorized.

Critical thinking: The ability to solve problems by utilizing skills, abilities, values, generalizations, etc., that were learned or memorized in other contexts.

Memory questions: Questions calling for recall of facts, generalizations, value judgements, definitions or other information previously learned by the student.

Translation questions: Questions requiring students to change previously learned materials into different language, symbols or forms.

Interpretation questions: Questions calling for skill in reading charts, maps, graphs, and in drawing inferences from statistics. The student must discover relationships from the materials being studied.

Application questions: The student must identify the problem and select from his repertoire of skills and knowledge the appropriate tools to apply in solving the problem.

Analysis questions: Questions calling for a conscious knowledge of the parts and forms of thinking used by the student in answering the questions.

Synthesis questions: Questions requiring students to apply knowledge in ways that are new to himself. Similar to application questions except that the skills and knowledge are applied in creative ways.

Evaluation questions: The student is asked to form value judgements based upon standards he establishes rather than those imposed by teachers.

Phase I or Baseline Phase: The period of time from March 1, 1971 to April 1, 1971, during which teachers were observed in their classrooms and test procedures to establish the percentage of higher order questions employed.

Phase II: The period of time from April 1, 1971 (approximately), to June 1, 1971, during which teachers were observed to determine the effectiveness of the inservice training.

Inservice training: A two-hour presentation of a condensation of various levels of questions, their uses and objectives, that was presented to teachers between observational Phases I and II of the study.

Observers: Six college students who were trained to rate the questioning behavior of teachers in the classroom and on tests.

METHOD

Twelve volunteer elementary school teachers served as subjects. Their classroom and test questioning behavior was monitored by members of a team of six college seniors who had been trained to categorize questions as either "memory level" or "higher order" using an eight category instrument adapted from Pate and Bremer's (14) "Questions Analyzer".

Include Table 1 About Here

The Questions Analyzer's eight categories were divided into four memory level (simple recall-one item; simple recall-multiple items; skills-demonstration; and skills-verbal) and four higher order question levels (example-singular; example-multiple; principle involved, and concept analysis).

Frequency tallies in the various categories were obtained during each observational period. Tests were analyzed by classifying questions as either "memory" or "higher order". These ratings were performed by four of the Observers who classified classroom questions.

Observers were selected from advanced undergraduate education majors at the University of Tennessee. Observer training was conducted in three phases. The first phase consisted of learning Sanders' (16) taxonomy of question levels and the definitions of the categories in the "Questions Analyzer". The Experimenter explained to Observers that a technique commonly known as "trouble shooting" could be employed to aid in classifying classroom questions. In this technique the Observer silently asks himself three questions, each of which divides the rating categories into successively smaller halves. An example of this technique follows:

Teacher: "What does 'compact' mean?"

(Observer must wait for student responses before rating the question since often the teacher's acceptance of replies aids in classificiation of questions.)

Student: "A Volkswagen is a 'compact'!"

(Teacher was referring to the Mayflower Compact, but she accepts this "answer".)

Observer (O hereafter) first asks, "was this a higher order or memory question?" (The answer to this question eliminates half the categories on the "Questions Analyzer" from further consideration.) Since the student's answer of an unusual sort was accepted--even though it might be argued that a Volkswagen is, in fact, a 'compact'--O decides that this was not a memory question. The second question O silently asks is, "Does the question call for an example or not?" (Had it been classified as a memory question O would have asked, "Does the

question ask for simple recall or not?") The answer here is, "yes". This divides the "Question Analyzer's" higher order categories and eliminates all possible categories save "example-singular" and "example-multiple". The next question O asks is, "One or more examples?" The answer is, "one". The question is categorized as "example-singular".

The example used was an actual problem encountered by one Observer who classified the questio as reported here.

The second phase of Observer training consisted of using a videotape of a random series of questions representing the eight levels of questions. These questions were independently rated by the six Observers. After several trials, a reliability coefficient of .985 was achieved using the technique known as Hoyt's Analysis of Variance (8).

The third phase of training consisted of rating questioning behavior of classroom teachers from video-tapes made in school settings. An inter-rater reliability coefficient of .90 was set as the standard for this phase of training. All Observers exceeded this standard on first exposure to the tapes, and therefore training was discontinued.

One Observer per classroom was used except for times when inter-rater reliability was checked during the experiment. The range of these checks (15 were made during the 16 weeks of Phases I and II) was from .84 to .99 with a mean of .96. Observers were instructed to sit near the back of the classrooms and to provide as little disruption of school routine as possible. Observers were rotated and never observed the same teacher on successive days. All

Observers visited all teachers and collected a total of 192 classroom observations. They also collected copies of teacher-made tests once each week producing a total of 192 tests.* Four of the Observers were used to analyze test questions.

The design of the study was as follows: The 12 volunteer teachers were randomly assigned to three groups immediately after they were selected, and prior to the Experimenter's personal acquaintance with any teacher other than having met with them during the explanation of the requirements of teachers in the proposed study. The three groups to which teachers were assigned were designated as the "Control", the "Inservice", and the "Feedback" groups.

During the first, or Baseline, phase of the study, all teachers were to provide the Experimenter with one copy of a teacher-made test each week for eight weeks and to permit at least one, and not more than two, 30 minute classroom sessions to be observed until a total of eight observations was achieved. The classroom observations were arranged from the teacher's schedule of classes and observing was scheduled during times when teacher-student discussion would take place. Teachers did not know when they would be observed and efforts were made to vary the time and day of observing. The procedure described was applied to all teachers. They were not told to which group they had been assigned during the Baseline Phase. The time period for Phase I observations was from March 1, 1971 to April 1, 1971. During this period 96 classroom observations were made and 96 tests were collected.

---

*As indicated in the example of rating classroom questions, the student's answer is helpful in categorizing a question. Since the tests collected did not contain accepted student answers, it was difficult to obtain an agreement between Observer ratings of test questions. It was found that Observers could agree upon the gross classification of either "memory" or "higher order" with the test question alone serving as a cue and the E was forced to adopt this classification method.

After completion of the Baseline observations, teachers in the Feedback and Inservice groups participated in a training program conducted by the Experimenter. This training program consisted of instruction in preparing classroom and test questions in accordance with materials presented in Sanders' (16) book. The text of this material was given to each teacher in the two groups for further study after the training program. The training program was completed in a two hour period on April 1, 1971.

Phase II of the study consisted of observing eight additional 30 minute classroom sessions for each teacher and collecting eight more teacher-made tests. Classroom observations were made once or no more than twice, weekly, and tests were collected weekly from each teacher by Observers or the Experimenter.

During the second phase of observations, the members of one group of four teachers, the Feedback group, were individually informed by the Experimenter of their performance on both teacher-made tests and classroom questioning behavior. The Experimenter revealed to the teachers the percentage of higher order questions and offered suggestions relating to utilizing materials provided during Inservice. In a few instances, the Experimenter suggested to members of the Feedback group that they restrict the use of higher order classroom questions because a factual basis had not been established from which their students could synthesize and evaluate related applications of the materials being studied.

RESULTS

Analysis of the data consisted of computing the percentage of higher order to total questions asked by each teacher in each group. Percentage of higher order questions were calculated for each phase of the study and for the gain score

between phases. These figures (Table 2) were then used to compare the three groups' performance during the various phases of study. Statistical tools employed were the Kruskal-Wallis analysis of variance by ranks; with a follow-up analysis using the Mann-Whitney U Test; and the Spearman rho ranked coefficient of correlation. All of these techniques are described in Siegel (18).

Include Table 2 About Here

Kruskal-Wallis comparisons of the average percentages of higher order questions indicated only chance differences among the groups in classroom questioning during both observational phases and in gain scores. Test questions differed significantly (p .049) during Phase II observations and subsequent analysis with the Mann-Whitney U Test showed that while both treatment groups differed significantly (p .029) with the control the difference the two treatment groups was not significant.

When the average percentages for test and classroom higher order questions were ranked and correlated it was found that none of the groups had sufficiently high coefficients for significance during Phase I (Control——.40; Inservice .20; and Feedback ——.80). However during Phase II both treatment groups achieve perfect positive coefficients of correlation (p .05) and the Inservice group also had a perfect positive correlation coefficient in its gain score rankings (p .05).

Because of the inability to collect test questioning data in the eight question categories, it was only possible to compare the two kinds of questions used by teachers by using the gross classifications of "memory" and "higher order". However average percentages of classroom questions in each category

were computed and comparisons made between the three groups across both observational phases and gain scores. In order to give the reader a general picture of how the groups compared, the percentages contributed by each group member were averaged for each question category during Phases I and II and for gain scores. These"collapsed" data were used to compile Table 3.

Table 3 Goes About Here

The percentage of each group member was ranked and the groups compared using the Kruskal-Wallis technique. The results of this analysis appear in Table 4.

Table 4 Goes About Here

The Phase II difference in the Example-singular category was further analyzed with the Mann-Whitney U Test and significant differences were noted (p .029) when each of the two treatment groups was compared with the control, but only a chance difference was noted when the two treatment groups were compared.

The Mann-Whitney U Test was also used to compare the groups found to differ significantly in gain score comparisons. Simple recall-singular showed significant differences (p .029) between each treatment group and the control but nonsignificant differences between the treatment groups. Simple recall-multiple showed a significant difference between the two treatment groups (p .014) and Example-singular revealed that the major portion of the variance among the groups resided in a significant difference (p .014) between the Feedback and Control groups.

DISCUSSION

The results reported in this study were consistent with the findings of other research (14, 20, 9, 5, 15). Teachers were observed to use a very high percentage of memory level questions with a particular emphasis on the factual recall question, and few teachers used more than a scattered number of higher order questions. Also consistent with the literature (15) was the reported finding of the present study that teachers use even fewer higher order test questions than they use in classroom presentations.

The consistency between classroom and test questioning behaviors was found to be either low or non-existant during Phase I comparisons. The reason for the inconsistency was because teachers who used a relatively high percentage of higher order classroom questions used a correspondingly low percentage of higher order test questions.

The apparent decrease in correlation between classroom and test higher order questioning during Phase II observations was an artifact of the statistical technique used. The large number of ties at 0 percent (especially in Phase I test questions) had the effect of improving the rankings of all of these scores, thereby reducing their differences from the classroom rankings with which the relationship is drawn. These ties (of 0 percent) disappeared once teachers began to use higher order questions during Phase II observations, and a more discriminating ranking showed realistic correlations between the two variables.

In the statistical analysis of data collected during Phase I, no significant differences were found in the correlation coefficients between classroom and test questioning behavior of the groups, and no significant differences in the higher order classroom questioning behavior were noted. No significant differences were

observed in the higher order test questioning behavior of the groups, and no significant differences in the eight categories of classroom questioning behaviors of the three groups were found. These findings indicate that, statistically, the questioning behavior of the three groups was the same during Phase I except for chance variability. While no specific hypotheses related to typical behavior of teachers were formulated, it was necessary to review the literature and establish a baseline of the typical questioning behavior of teachers who participated in the present research. This was done to establish that teachers in the present study were not unusual in their questioning behavior and to provide a baseline from which differences in the two training groups could be measured.

In summary, the results of analysis of the data collected during Phase I produced data similar to other research concerned with the questioning practices of teachers. While not supported by research, the common complaint by educators and students that teachers do not "test the same way they teach" was validated by the low correlation coefficients found between classroom and test questioning during Phase I.

Although previous studies (13, 4, 17) reported that it is possible to alter teachers' higher order questioning behavior, it was felt that the weakness of the designs of these studies reduced the confidence one could place in the findings. For this reason, the present study incorporated a longer time period, more observations, and measured two aspects (classroom and tests) of the questioning behavior of teachers. The findings of this study were that teacher questioning behavior can be changed to use fewer factual recall type questions and more higher order questions. Teachers can also be trained to use more variety in the levels of

questions they ask, rather than depending on one kind (usually recall of specific information). These findings were in accord with previous reports (10, 17, 4, 2).

The teachers participating in this study were aware of their own weaknesses in ability to prepare higher order questions, and, during Phase II observations, all teachers made comments about how pleased they were to be participating in a study specifically aimed at this problem. This positive attitude possibly contributed to the increased use of higher order questions during the post-training observations.

A certain amount of increased use of higher order questions by teachers in the Control group was noted during Phase II. Part of this increase could be attributed to Hawthorne Effect, or possibly Control group members were informed by other group members about the training they had received.

Since most of the increased higher order questioning of the Control group was contributed by C_1 and C_4 (who work as a "team" in the same school), it was the feeling of the Experimenter that the "leak" (if there was one) applied only to them. Investigation disclosed that a student teacher from a nearby college had been assigned to the two teachers involved shortly before the increased use of higher order questions began to manifest itself. Possibly this student teacher, who had been trained in higher order questioning techniques, was unknowingly influencing her two supervising teachers.

The reports by Meehan (12), Claus (2), Konetski (10), and Manson (11) stating that while training teachers specifically to use higher order questions caused increases over teachers not specifically trained, but that training methods do not produce differing behavior among themselves, were partially confirmed in this study. In many cases, where significant differences were

observed among the groups, the major portion of the difference could be traced to differences between the Feedback and Control, and Inservice and Control group, but not between the Feedback and Inservice groups.

An important differentiation should be made between studies reporting the effects of feedback on higher order questioning behavior. Some of the studies provided feedback on performance alone (11, 13) while others involved feedback of performance and also a training program (2, 10, 12). The present study is similar to the designs of the latter group cited in that it included a specific training program as well as reports of performance. Claus (2) and Manson (11) both reported that feedback, by itself, was not a strong enough stimulus to produce a significantly different result over those attributed to a training program alone. The findings of this study were generally in accord with these reports. Possibly further research related to specific feedback variables is indicated since the findings of the researchers (13, 11, for examples) are diametrically opposed.

The improvement in positive correlations between teacher's higher order classroom and test questions in both the Inservice and Feedback groups, and the lack of difference in correlation in the Control group, is evidence that teachers can be trained to "test as they teach". The fact is that few teachers (including those in this study, during Phase I) actually do ask the same types of questions on tests that they ask in the classroom. The puzzling aspect of this situation is that research has ignored the problem. Hopefully the findings in this study will serve as a beginning to further research into this important area of educational practice.

The statistical analysis of data collected during Phase II provides evidence that the correlation between higher order classroom and test questioning behavior

can be increased. Remarkable improvement occurred in both the Feedback and Inservice groups (from —.80 during Phase I to +1.00 ($p < .05$) in Phase II) for the Feedback group, and from +.20 to +1.00 in Phase II for the Inservice group). No significant change occurred in theControl group, indicating that training has a differential effect in improving the correlation between classroom and test questioning practices of teachers.

Since some significant increases in higher order classroom and test questioning were noted during Phase II observations and the groups were found to be statistically equal during Phase I, a rather tentative acceptance of Hypothesis II is indicated. Added support for this acceptance is provided by significant gain score differences between Phases I and II, as well as the indication of trends toward significance in Phase II and gain score observations.

In summary, the findings of this study were that teachers' questioning behavior includes few higher order classroom questions and even fewer higher order test questions. These findings are consistent with the reports of Pate and Bremer (14) and Pfeiffer and Davis (15).

The training to which teachers were exposed in the present study had the effect of increasing the number of higher order questions employed. Studies by Mittelstadt (13), Cunningham (4), and Schreiber (17) reported similar results. The reports by Meehan (12), Claus (2), Konetski (10), and Manson (11) that feedback is not sufficiently strong, in itself, to alter higher order questioning behavior of teachers were upheld. Findings of the present study, while not supported by other research, indicate that it is possible to train teachers to bring classroom teaching and test questioning behaviors into a higher degree of correlation.

TABLE I

QUESTION CATEGORIES: (from "Questions Analyzer")

Convergent questions:

(a) Simple recall - one item: Remembering a specific fact.

(b) Simple recall - multiple: Choosing the correct answer from among several possibilities.

(c) Skills - demonstration: Requires student to exhibit mastery of knowledge by performing certain acts.

(d) Skills - verbal: Requires student to explain in his own words the processes involved in performing certain acts.

Divergent questions:

(a) Example - singular: Student provides one illustrative sample of an idea under discussion.

(b) Example - multiple: Student provides several illustrative samples of an idea under discussion.

(c) Principle involved: Teacher provides an example illustrating an idea and student provides the related idea.

(d) Concept analysis: Student is asked to evaluate, establish standards, infer relationships, and create new applications of known facts.

TABLE 2

AVERAGE PERCENTAGES OF HIGHER ORDER QUESTIONS FOR EACH TEACHER DURING PHASES I AND II AND GAIN SCORES FOR CLASSROOM AND TEST QUESTIONS

	Classroom Questions			Test Questions		
Ss	Phase I	Phase II	Gain	Phase I	Phase II	Gain
Control						
C1	11.0	14.8	3.8	0.7	2.6	1.9
C2	11.1	12.8	1.7	10.5	6.6	-3.9
C3	6.1	5.5	-0.6	2.0	0.9	-1.1
C4	12.3	22.2	9.9	0.0	4.7	4.7
Avg. Percent	10.0	12.4	0.6	3.0	3.4	0.4
Inservice						
I1	8.9	18.3	9.4	0.0	9.0	9.0
I2	6.5	35.1	28.6	0.7	30.0	29.3
I3	15.7	16.4	0.7	16.7	6.5	-10.2
I4	6.7	46 6	39.9	0.6	45.0	44.4
Avg. Percent	9.6	28.4	18.8	5.2	16.1	10.9
Feedback						
F1	3.2	8.1	4.9	4.5	6.4	1.9
F2	6.0	36.3	30.3	0.6	52.6	52.0
F3	12.9	25.7	12.8	1.3	26.0	24.7
F4	5.9	23.1	17.2	2.3	24.0	21.7
Avg. Percent	6.2	22.2	16.0	2.0	29.0	27.0

TABLE 3
*

AVERAGE PERCENTAGE OF CLASSROOM QUESTIONS
IN EACH QUESTION CATEGORY PER GROUP

	Memory Level				Higher Order			
Groups	SR (1)	SR (M)	S(D)	S(V)	E(1)	E(M)	PI	CA
				PHASE I				
Control	57.9	7.6	14.2	13.3	3.7	0.3	0.4	2.4
Inservice	71.8	7.8	6.2	5.5	2.7	0.7	2.8	2.1
Feedback	60.0	8.8	16.0	10.0	4.0	0.1	1.0	0.2
				PHASE II				
Control	54.6	7.6	10.6	14.0	2.8	2.0	2.4	3.6
Inservice	45.9	11.7	5.7	9.5	6.8	3.8	6.2	10.4
Feedback	38.9	7.3	15.0	16.0	7.4	2.9	5.1	7.1
				GAIN SCORES				
Control	-3.4	0.0	-8.7	7.9	-4.1	2.0	2.3	1.8
Inservice	-26.3	3.8	1.1	4.2	4.4	2.7	3.3	8 2
Feedback	-17.9	-1.6	-0.6	6.5	2.9	2.8	4.4	7.2

* These data were prepared by averaging the percentage of total questions in each category for each member of each group. This accounts for the non-additivity of the data.

TABLE 4

KRUSKAL-WALLIS ANALYSIS OF AVERAGE PERCENTAGE
OF CLASSROOM QUESTIONS IN EACH CATEGORY

	Memory Level				Higher Order			
Phases	SR (1)	SR (M)	S(D)	S(V)	E(1)	E(M)	PI	CA
I	4.20	0.77	2.96	2.62	0.31	1.23	5.39	5.61
II	5.16	3.27	5.39	0.85	5.82*	1.23	2.92	2.85
Gain Scores	5.85*	5.71*	0.38	1.29	5.70*	0.38	0.44	2.97

* p .049

References

1. Bloom, B. S. (ed.) Taxonomy of Educational Goals: Handbook I. Cognitive Domain. New York: David McKay Co., Inc., 1956.

2. Claus, K. E. S. The Affects of Modeling and Feedback Variables On Higher-Order Questioning Skills. Doctoral dissertation, Stanford University. Ann Arbor, Mich.: University Microfilms, 1968, No. 69-207.

3. Cornell, E. F. Observational Learning of Question-Asking Behavior Through the Medium of Audio-Tapes. Doctoral dissertation, University of Maryland. Ann Arbor, Mich.: University Microfilms, 1969, No. 70-15, 394.

4. Cunningham, R. T. A Descriptive Study Determining the Effects of Instruction Designed to Improve the Question-Phrasing Practices of Prospective Elementary Teachers. Doctoral dissertation, Indiana University. Ann Arbor, Mich.: University Microfilms, 1968, No. 69-6729.

5. Davis, O. L., and D. C. Tinsley. Cognitive Objectives Revealed by Classroom Questions Asked by Social Studies Student Teachers. Peabody Journal of Education, 1967, 45, 21-26.

6. Fenton, E. The New Social Studies. New York: Holt, Rinehart, and Winston, Inc., 1967.

7. Guilford, J. P. Intelligence, Creativity, and Their Educational Implications. San Diego: R. R. Knapp, 1968.

8. Hoyt, C. J. Test Reliability Estimated by Analysis of Variance. Psychometrika, 1941, 6, 153-160.

9. Hunkins, F. P. The Influence of Analysis and Evaluation Questions on Achievement in Sixth Grade Social Studies. Educational Leadership Research Supplement, 1968, 1, 326-332.

10. Konetski, L. C. Instructional Effect on Questions Asked by Preservice Science Teachers. Doctoral dissertation, Indiana University. Ann Arbor, Mich.: University Microfilms, 1969, No. 70-7974.

11. Manson, G. A. The Effect of Immediate and Postponed Observer Feedback on the Acquisition of Higher Order Questioning Skills by Prospective Teachers. Doctoral dissertation, University of Washington. Ann Arbor, Mich.: University Microfilms, 1970, No. 71-1000.

12. Meehan, Sister M. T. The Effects of Instruction Based on Elements of Critical Reading Upon the Questioning Patterns of Preservice Teachers. Doctoral dissertation, Indiana University. Ann Arbor, Mich.: University Microfilms, 1970, No. 70-26, 940.

13. Mittelstadt, W. Description and Analysis of Questioning Strategies of Elementary Education Student Teachers Prior to and Following the Implementation of an Inquiry Observation Instrument. Doctoral dissertation, Wayne State University. Ann Arbor, Mich.: University Microfilms, 1969, No. 70-3432.

14. Pate, R. T. and N. H. Bremer. Guiding Learning Through Skillful Questioning. Elementary School Journal, 1967, 67, 417-422.

15. Pfeiffer, I., and O. L. Davis, Jr. Teacher-Made Examinations: What Kind of Thinking Do They Demand? NASSP Bulletin 49, 1965, pp. 1-10.

16. Sanders, N. M. Classroom Questions: What Kinds? New York: Harper and Row, 1966.

17. Schreiber, J. E. Teachers' Question-Asking Techniques in Social Studies. Doctoral dissertation, University of Iowa. Ann Arbor, Mich.: University Microfilms, 1967, No. 67-9099.

18. Siegel, S. Nonparametric Statistics for the Behavioral Sciences. New York: McGraw-Hill Book Company, 1956.

19. Torrance, E. P. Education and the Creative Potential. Minneapolis: University of Minnesota Press, 1963.

20. Torrance, E. P. Creativity in the Classroom. Instructor, 1964, 74, 35.